4.3 Future QA/QC Efforts and Tools

Quality Assurance and Quality Control (QA/QC) have always been a vital part of GIS work because they allow users to ensure a level of accuracy and quality for all published datasets. Standardizing QC efforts is equally important, because otherwise data could be made available that needs to be QC'd prior to its publication. Handling these datasets can be burdensome without a standardized process for handling specific types of data. Hence, efficient standardized QC processes will be developed and incorporated into the maintenance program at OCTO GIS.

While source agencies will have varying levels of agency QC standards, having published core standards for data distribution and storage permits appropriate QA/QC tools to be developed and incorporated into customized ArcGIS map documents. OCTO GIS has already developed several programs that ensure the data meets several of the standards listed in section five. Moreover, exploiting ArcSDE attribute domains, subtypes, and topology will further check and standardize data. OCTO GIS intends to work with agencies to license and/or develop tools to help spatial data editors measure and improve adherence to quality goals.

Types of checks that can be automated include:

Single Column Attribute Tests

Unique Values

Detects non-unique values for a column. Useful for columns that are being used as primary keys or columns that represent feature identifiers where uniqueness is required.

Coded Domains

Detects invalid values based on pre-established geodatabase Coded Domains, existing DBMS tables or a user-defined list of valid values. All values for a tested column must contain only valid values.

Coded Ranges

Detects out-of-range values based on pre-established geodatabase Range Domains or user-defined minimums and maximums. All values for a tested column must contain only valid values.

Single Values

Detects values other than the specified single value for a column. All values for a column must equal the specified value.

NULL Values

Detects NULL, blank and/or zero values for a column. Can detect NULL occurrences if NOT NULL constraints are disabled while loading the data into the geodatabase. Also detects blank and/or zero values occurring in legacy INFO tables.

Non – Standard Values

Detects non-standard keyboard values such as !@# and ? in a column. A list of these values may be configured for each column.

Column Format

Detects values that do not match the specified format. May set format to uppercase, lowercase, numeric, or non-numeric.

Column Length

Detects values that are not the specified length.

Multiple Column Attribute Tests

Custom SQL Query

Reports results of a user-defined query to the database. Read-only queries are allowed.

Multiple Column Unique

Detects non-unique values for multiple columns. This test is useful when searching for non-unique Book/Page/Parcel combinations.

Referential Integrity Test

General Table Relationship

Detects unmatched relationship keys between two tables in a join. No orphan table rows are allowed.

Spatial Relationship Test

Distance

Calculates the distance from geometries in one feature class to geometries in another. A distance tolerance may be used. Additionally, attribute filters may be applied to either feature class. A spatial filter may be applied to the test to limit the scope of the features tested. This test detects features that cannot participate in a geometric network because they lack coincidence.

In addition to the standardized QC documentation and software, custom QC tools must be developed and incorporated into the regular maintenance procedures. One of those tools will be the new OCTO GIS website. One aspect that will assist in future QC efforts at OCTO will be the new error tracking system, called Error Reporting and Review Online Registry System (ERRORS). It will be a portion of the website that focuses on issue monitoring. Internal and external users of the site will be able to:

- Report errors that they find in the data.

- Track each error issue's status as it is being resolved; i.e., Status = Error Submitted, Status = Error under Review, Status = Error Pending, Status = Error Resolved (see note for distribution instructions), etc.
- Get notification that their dataset has been updated and is published or in the queue.

Having this portion of the new website in place will help various agencies and users that use OCTO GIS's datasets to communicate more easily, as well as stay informed about datasets that are made available through OCTO GIS. On the other hand, this system will also handle agency and public feedback on datasets. The DBA will run daily statistics on the datasets in order to generate a progress report so that staff and data maintenance can be monitored. The report will list who is currently creating features, how many features they have created, what project they are working on, and how many hours they have been working on data creation for the project. More sections of this report will be generated when additional needs arise.